Advances in Applied Mathematics **22,** 119–153 (1999)
Article ID aama.1998.0625, available online at http://www.idealibrary.com on IDEAL®

A Convergence Speeding Algorithm with Applications to Numerical Integration

John Stalker*

Received January 1, 1998; accepted September 19, 1998

A new algorithm is presented for accelerating the convergence of sequences possessing an asymptotic expansion. This method is compared to methods already shown. Explicit error estimates are given and the algorithm is shown to be nearly optimal. The algorithm is applied to the problem of numerical integration, and is shown to give good results for real analytic integrands. © 1999 Academic Press

1. INTRODUCTION

The situation considered in this paper is as follows. A sequence a_n is known to possess an asymptotic expansion; i.e., there are sequences c_k, called *coefficients*, and e_m, called *error constants*, such that

$$\left| a_n - \sum_{k=0}^{m-1} c_k n^{-k} \right| \leq e_m n^{-m}. \tag{1}$$

One wishes to compute $c_0 = \lim_{n\to\infty} a_n$ to within specified accuracy from the a_n and some estimates on the growth of the e_m.

The purpose of this paper is to present and to analyze a fast algorithm for this problem. If the work required to compute a_n to d digits is polynomial in n and d and the e_m grow at worst factorially in m, this algorithm computes c_0 in time polynomial in the number of digits desired. Error estimates are given and the algorithm is shown to be "nearly optimal" in a sense made precise in Section 11. A brief description of the algorithm is given in the first two paragraphs of the next section.

The algorithm reduces, for special (nonoptimal) values of the parameters, to a variant of Richardson extrapolation, introduced in [13]. Even in

* Partially supported by NSF.

0196-8858/99 $30.00

this case we believe the error analysis given here is essentially new. In general the algorithm given here outperforms methods based on Richardson extrapolation. The algorithm also outperforms those of Bulirsch–Stoer type [3].

The motivating example is that of numerical integration. Let f be a smooth function on $[0,1]$ and let a_n be the Riemann sum corresponding to a regular mesh of size n^{-1}. The Euler–MacLaurin summation formula provides estimates of the form (1). The number of evaluations of f required is linear in n. Without further regularity assumptions nothing can be said about the e_m. If f is, for example, real analytic then

$$e_m \leq Cm!(\pi r)^{-m},$$

with some constant C, provided r is chosen smaller than the minima of the radii of convergence at points of $[0,1]$. In this case, our algorithm gives a method for computing $\int_0^1 f(t)\,dt$ with a number of function evaluations which grows only quadratically in the number of digits desired.

Returning to the general problem, we see from (1) that

$$|c_0 - a_n| \leq \epsilon,$$

if $n \geq e_1/\epsilon$. While this gives a conceptually very simple algorithm for computing c_0 it is not usually very efficient. The earliest systematic improvement is due to Aitken who defined in [1] an auxiliary sequence,

$$A_m = 2a_{2m} - a_m.$$

It is easy to see from (1) that

$$|c_0 - A_m| \leq \epsilon,$$

if

$$m \geq \sqrt{3e_2/2\epsilon}.$$

In the case of numerical integration this reduces to Simpson's rule, [4]. There are innumerable variations on this algorithm, most of which share two common features: they make no assumption on the rate of growth of e_m, and the n for which a_n must be computed grows exponentially in the number of digits desired. Unfortunately the latter feature follows inevitably from the former.

The first to escape from this trap in any systematic way was Romberg in [14]. He was able to produce an auxiliary sequence A_m, depending linearly on $a_{2^0}, a_{2^1}, \ldots, a_{2^{m-1}}$ with

$$|c_0 - A_m| \leq Ce_m 2^{[-m(m+1)]/2}.$$

Under very weak assumptions on e_m the convergence is very rapid. The defect is that 2^{m-1} is very large. For certain applications to special functions the time required to compute $a_{2^{m-1}}$ is linear in m, and the algorithm is very efficient. For numerical integration and many other applications the time is proportional to 2^{m-1} and the defect is serious indeed; the effort required to compute $\int_0^1 f(t)\,dt$ is exponential in the square root of the number of digits desired. This has not prevented the algorithm from enjoying a certain degree of popularity. It is, for example, highly recommended in [5].

Various algorithms have been suggested which, for restricted classes of functions, produce faster convergence than that of Romberg. The theoretical investigation of such algorithms does not seem to have proceeded very far. We have not been able to locate, for any of these algorithms, either a description of the class of functions to which it is applicable or any statement concerning its rate of convergence.

2. OUTLINE

It is impossible to give a complete description of our algorithm in a paragraph or two, but it is perhaps possible to convey the general idea. The hypothesis (1) says that a_n is well represented as a polynomial,

$$\sum_{k=0}^{m-1} c_k n^{-k},$$

of degree $m-1$ in $1/n$ plus a small error. Multiplying by n^m gives a polynomial,

$$\sum_{k=0}^{m-1} c_k n^{m-k},$$

of degree m in n. This operation is, of course, represented by a diagonal matrix on the vector space of sequences with the obvious basis. Note that the coefficient of the constant term in the old polynomial becomes the coefficient of the leading term in the new polynomial. It is this coefficient which we seek to determine. Applying a difference operator m times will annihilate all but this term. The matrix which computes successive powers of the difference operator is very simple. It looks like Pascal's triangle except with alternating signs. If the error in approximating a_n by a polynomial were zero then we would have succeeded in writing c_0 as a linear combination of the elements $a_1, \dots, a_m$ with coefficients $w^1_{m,1}, \dots, w^1_{m,m}$ which are given as an alternating sign times a binomial coefficient times n^m. Of course the error is not, in general, zero and we must estimate the difference between c_0 and this linear combination. This is not hard and we are able to show that the errors tend rapidly to zero under fairly general hypotheses.[1]

This argument can be improved in several ways. If the coefficients of odd powers of $1/n$ are known to be zero we should expect to be able to use this information. This can be done by using centered differences in place of one-sided differences. The arguments are sufficiently similar that we handle both cases at once, using a superscript to distinguish them. We can also make the algorithm converge for a wider class of sequences, at the expense of making the approximants harder to compute, by restricting attention to a subsequence of the form a_{dn^r}. The most interesting improvement however, and technically the most demanding, is to exploit the alternation of signs. This suggests that smoothing operators produce cancellation. This is indeed the case. The successive powers of a smoothing operator are easily computed. The matrix again looks, except for powers of 2, like Pascal's triangle, this time without the alternating sign. We again obtain approximations to c_0 given by linear combinations of $a_1, \dots, a_m$ with coefficients expressible in terms of binomial coefficients and powers of n, but the error analysis is much more difficult. This analysis reveals that the use of smoothing operators produces a significant improvement both in speed and generality.

The organization of the paper is as follows. Section 3 contains some elementary observations about difference operators required in the succeeding section. Section 4 presents the promised algorithm for extrapolation of sequences, though only in a somewhat primitive form. This is however already sufficient with many interesting applications. Section 5 reviews the Euler−MacLaurin summation, which the succeeding section requires in a form slightly different from that usually given. Section 6 applies the extrapolation algorithms of Section 4 to give fast algorithms for numerical integration. Section 7 considers the small adjustments needed to take into account finite accuracy in the computation of the a_n. It is clear that, in order to get d digits of accuracy in the final result, we will need to

[1] It is also possible to view this as an interpolation problem. In fact this is the approach taken by Richardson and by Bulirsch and Stoer. The formulation in terms of difference operators given in the preceding text, besides simplifying the error analysis, has the advantage that it suggests the use of discrete smoothing operators, as described in the following paragraph.

know the original sequence to $O(d)$ digits of accuracy. The point of Section 7 is that we need only $O(d)$ digits. The reader is at this point in possession of some fast algorithms and completely explicit error estimates and may wish to stop here.

The remaining sections are more technically demanding. Section 8 contains estimates on certain sums needed in the error analysis of an improved form of the algorithm of Section 4. This improvement depends on the properties of certain discrete smoothing operators whose theory is developed in Section 9. The improved algorithm is then presented and applied to numerical integration in Section 10. The question naturally arises of the extent to which further improvement is possible. Section 11 shows that this is rather limited. The integrals treated in the preceding sections, one variable integrals without singularities on finite intervals, are intended primarily to illustrate the general technique without excessive complication. The same technique, with certain modifications, is applicable to the more general integrals which arise in practice, and Section 12 contains some suggestions for any readers who need to pursue this further.

3. DIFFERENCE OPERATORS AND INTERPOLATION

In this section we make precise the idea, mentioned in the previous section, of using difference operators as a tool for extrapolation of sequences. We define two difference operators D_1 and D_2 on $\mathbf{C}[x]$ by

$$(D_1 p)(x) = p(x+1) - p(x),$$

and

$$(D_2 p)(x) = p\left(x + \tfrac{1}{2}\right) - p\left(x - \tfrac{1}{2}\right).$$

The following fomulae are all easily proved by induction on m,

$$(D_1^m p)(x) = \sum_{n=0}^{m} (-1)^{m+n} \binom{m}{n} p(x+n),$$

$$\left(D_2^{2m} p\right)(x) = \sum_{n=-m}^{m} (-1)^{m+n} \binom{2m}{m+n} p(x+n),$$

$$D_l^m x^k = \begin{cases} 0, & \text{if } m > k, \\ \dfrac{k!}{(k-m)!} x^{k-m} + \text{lower order terms}, & \text{if } m \le k. \end{cases}$$

We can use the first two of these to rewrite the last as

$$\sum_{n=1}^{m}(-1)^{m+n}\binom{m}{n}n^{k}=\begin{cases}0, & \text{if } m>k,\\ m!, & \text{if } m=k,\\ ?, & \text{if } m<k,\end{cases}$$

$$\sum_{n=1}^{m}(-1)^{m+n}\binom{2m}{m+n}n^{2k}=\begin{cases}0, & \text{if } m>k,\\ \dfrac{(2m)!}{2}, & \text{if } m=k,\\ ?, & \text{if } m<k.\end{cases}$$

The question marks indicated complicated expressions which are irrelevant to our present purposes. Our extrapolation algorithm will construct an auxiliary sequence as a weighted sum of elements of an original sequence, with weights $w_{m,n}^{l}$ given by

$$\begin{aligned} w_{m,n}^{1}&=\frac{(-1)^{m+n}}{m!}\binom{m}{n}n^{m},\\ w_{m,n}^{2}&=(-1)^{m+n}\frac{2}{(2m)!}\binom{2m}{m+n}n^{2m}.\end{aligned} \tag{2}$$

With this notation and $l=1$ or 2 the preceding formulae take the form,

$$\sum_{n=1}^{m}w_{m,n}^{l}n^{-lk}=\begin{cases}1, & \text{if } k=0,\\ 0, & \text{if } 0<k<m,\\ ?, & \text{otherwise}.\end{cases}$$

This formula is required for the proof of

LEMMA 1. *Let $l=1$ or 2. If*

$$p_n=\sum_{k=0}^{m-1}c_k n^{-lk},$$

and

$$P_m=\sum_{n=1}^{m}w_{m,n}^{l}p_n,$$

with $w_{m,n}^{l}$ given by (2), *then $P_m=c_0$.*

Proof. Then we have

$$\begin{aligned} P_m&=\sum_{n=1}^{m}w_{m,n}^{l}\sum_{k=0}^{m-1}c_k n^{-lk},\\ &=\sum_{k=0}^{m-1}c_k\sum_{n=1}^{m}w_{m,n}^{l}n^{-lk},\\ &=c_0.\end{aligned}$$

We want to apply our difference operators to sequences of the form polynomial plus small error. We have just seen that these operators pick out the leading term of the polynomial. We must now check that they do not magnify the error unacceptably.

LEMMA 2. *Let $l = 1$ or 2. If*

$$|r_n| \leq e_m n^{-lm},$$

and

$$R_m = \sum_{n=1}^{m} w_{m,n}^l r_n,$$

with $w_{m,n}^l$ given by (2), *then*

$$|R_m| \leq \frac{2^{lm}}{(lm)!} e_m.$$

Proof. We have

$$|R_m| \leq \max_{1 \leq n \leq m} |r_n n^{lm}| \sum_{n=1}^{m} |w_{m,n}^l n^{-lm}|.$$

The first term is bounded by e_m by hypothesis. The second is

$$\frac{1}{m!} \sum_{n=1}^{m} \binom{m}{n} \leq \frac{2^m}{m!},$$

if $l = 1$ or

$$\frac{2}{(2m)!} \sum_{n=1}^{m} \binom{2m}{m+n} \leq \frac{2^{2m}}{(2m)!},$$

if $l = 2$. In fact, because the $n = 0$ term is missing, we get a slightly better result,

$$|R_m| \leq (1 - 2^{-m}) \frac{2^{lm}}{(lm)!} e_m,$$

if $l = 1$ or

$$|R_m| \leq \left(1 - 2^{-2m}\binom{2m}{m}\right) \frac{2^{lm}}{(lm)!} e_m,$$

if $l = 2$.

4. SIMPLE EXTRAPOLATION ALGORITHMS

The following represents the simplest case of our algorithm. Nonetheless it is still sufficiently powerful for may interesting examples.

THEOREM 3. *Let $l = 1$ or 2. If*

$$\left| a_n - \sum_{k=0}^{m-1} c_k n^{-lk} \right| \leq e_m n^{-lm}, \tag{3}$$

and

$$A_m = \sum_{n=1}^{m} w_{m,n}^l a_n, \tag{4}$$

then

$$|c_0 - A_m| \leq \frac{2^{lm}}{(lm)!} e_m. \tag{5}$$

In particular if

$$e_m \leq C(lm)! p^{-lm}, \tag{6}$$

with $C > 0$ and $\rho > 2$ then A_m converges geometrically to the same limit c_0 as a_n.

Proof. We define

$$p_n = \sum_{k=0}^{m-1} c_k n^{-lk},$$

$$P_m = \sum_{n=1}^{m} w_{m,n}^l p_n,$$

$$r_n = a_n - p_n,$$

$$R_m = \sum_{n=1}^{m} w_{m,n}^l r_n.$$

Clearly $R_m = A_m - P_m$. Lemmas 1 and 2 give

$$P_m = c_0,$$

and

$$|R_m| \leq \frac{2^{lm}}{(lm)!} e_m.$$

If the estimate (6) holds then

$$|R_m| \leq C\left(\frac{\rho}{2}\right)^{-lm}.$$

As in Lemma 2, we can prove slightly sharper estimates,

$$|c_0 - A_m| \leq (1 - 2^{-m})\frac{2^{lm}}{(lm)!}e_m,$$

if $l = 1$ or

$$|c_0 - A_m| \leq \left(1 - 2^{-2m}\binom{2m}{m}\right)\frac{2^{lm}}{(lm)!}e_m,$$

if $l = 2$.

The following theorem shows how to modify the preceding algorithm when the hypothesis (6) is violated. It is based on the observation that, if a_n possesses an asymptotic expansion in n^{-1}, then so does the subsequence $a_{p(n)}$ for any polynomial p. The higher the degree of the polynomial the slower the growth of the error constants for this subsequence. We apply this observation to the polynomial $p(n) = dn^r$.

THEOREM 4. *Let $l = 1$ or 2. If*

$$\left|a_n - \sum_{k=0}^{m-1} c_k n^{-lk}\right| \leq e_m n^{-lm},$$

and

$$A_m = \sum_{n=1}^{rm} w_{rm,n}^{l} a_{dn^r},$$

then

$$|c_0 - A_m| \leq \frac{2^{lrm}}{d^{lm}(lrm)!}e_m.$$

In particular, if

$$e_m \leq C(lrm)!\rho^{-lrm}, \tag{7}$$

with $\rho > 2d^{-1/lm}$ then A_m converges geometrically to the same limit c_0 as a_n.

Proof. By hypothesis,

$$\left| a_{dn^r} - \sum_{k=0}^{m-1} c_k d^{-k} n^{-lrk} \right| \le e_m d^{-lm} n^{-lrm}.$$

Define

$$\begin{aligned} p_n &= \sum_{k=0}^{m-1} c_k d^{-lk} n^{-lrk}, \\ P_m &= \sum_{n=1}^{rm} w^l_{rm,n} p_n, \\ r_n &= a_{rn^l} - p_n, \\ R_m &= \sum_{n=1}^{rm} w^l_{rm,n} r_n. \end{aligned}$$

Clearly $R_m = A_m - P_m$. Lemmas 1 and 2 give

$$P_m = c_0,$$

and

$$|R_m| \le \frac{2^{lrm}}{d^{lm}(lrm)!} e_m.$$

If the estimate (7) holds then

$$|R_m| \le C\left(\frac{d\rho^r}{2^r}\right)^{-lm}.$$

5. EULER–MACLAURIN SUMMATION

If we are to apply the results of the preceding section to a sequence of Riemann sums then we need to know that these possess an asymptotic expansion, and we need some way to estimate the error constants in this expansion. Both are provided by the Euler–MacLaurin summation formula.

The original arguments of Euler [6] and MacLaurin [12] do not give explicit formulae for the error term. This makes them practically unusable. The first to give an explicit remainder was Jacobi [9]. Here we need the error term in a slightly unusual form. Because the result admits a short and elegant proof it seems advisable to give a self-contained proof, particularly because, as noted in Section 12, some readers may need extensions of the formula to other sequences of Riemann sums.

The proof is based on the properties of the Bernoulli polynomials B_k, or rather their periodic extensions B_k^*. We begin by reviewing these. The properties we need are most easily proven from their Fourier series. If $k > 0$ we define

$$B_k^*(x) = \begin{cases} (-1)^{(k+2)/2} 2\left[k!/(2\pi)^k\right] \sum_{n=1}^{\infty} \dfrac{\cos(2\pi nx)}{n^k}, & \text{if } k \text{ is even,} \\ (-1)^{(k+1)/2} 2\left[k!/(2\pi)^k\right] \sum_{n=1}^{\infty} \dfrac{\sin(2\pi nx)}{n^k} & \text{if } k \text{ is odd.} \end{cases} \tag{8}$$

Uniform convergence is clear for $k > 1$, and is false for $k = 1$. The following elementary argument shows, however, that B_1^* converges uniformly on any closed interval not containing an integer. Because the sum is clearly **Z**-periodic it suffices to consider only $x \in (0, 1)$. Making the convenient substitution $\xi = x - \frac{1}{2}$, the definition of the infinite sum is

$$B_1^*(x) = -\frac{1}{\pi} \lim_{N\to\infty} \sum_{n=1}^{N} (-1)^n n^{-1} \sin(2\pi n\xi).$$

After applying various trigonometric identities we find

$$B_1^*(x) = \lim_{N\to\infty} \int_0^{\xi} \left(1 - (-1)^N \frac{\cos(2N+1)\pi t}{\cos \pi t}\right) dt.$$

Splitting the integral and integrating the second term by parts gives

$$\begin{aligned} B_1^*(x) = \xi - &\lim_{N\to\infty} \frac{(-1)^N}{2N+1} \\ &\times \left(\frac{1}{\pi} \frac{\sin(2N+1)\pi\xi}{\cos \pi\xi} - \int_0^{\xi} \frac{\sin(2N+1)\pi t \sin \pi t}{\cos^2 \pi t}\, dt \right) \end{aligned}$$

The parenthesized expression is clearly bounded on any closed subinterval of $(0, 1)$, so

$$B_1^*(x) = x - \tfrac{1}{2}, \tag{9}$$

on such an interval, with uniform convergence.

This uniform convergence allows one to differentiate term by term,

$$\frac{d}{dx} B_k^*(x) = k B_{k-1}^*(x), \tag{10}$$

whenever $k > 2$ or $k = 2$ and $x \notin \mathbf{Z}$. If $k > 1$ then Cauchy's theorem on uniform limits of continuous functions gives

$$\lim_{\epsilon \to 0} B_k^*(\epsilon) = (-1)^k \lim B_k^*(1 - \epsilon) = B_k^*(0),$$

and the right-hand side is zero if k is odd. If $k = 1$ the second equality is false, as (9) gives

$$\lim_{\epsilon \to 0} B_1^*(\epsilon) = -\lim B_1^*(1 - \epsilon) = -\tfrac{1}{2}.$$

We will need various norms of the function $B_{2m}^* - B_{2m}^*(0)$. It can be shown that

$$\| B_{2m}^* - B_{2m}^*(0) \|_{L^p([0,1])} \le C_p\bigl(1 + \epsilon_p(m)\bigr) \frac{(2m)!}{(2\pi)^{2m}},$$

where

$$C_p = 4\pi^{(1-p)/2p} \Gamma\left(\frac{3p-1}{2p-2}\right)^{(p-1)/p} \Gamma\left(\frac{2p-1}{p-1}\right)^{(1-p)/p},$$

and $|\epsilon_p(m)|$ is bounded by $\zeta(2m) - 1$ and hence is asymptotic to zero. The proof requires only the triangle inequality for the L^p norm and the definite integrals of [7]. Only the cases $p = 1, 2, \infty$ arise with any frequency in practice. The values of C_p and ϵ_p in these cases are

$$\begin{aligned} C_1 &= 2, & \epsilon_1(m) &= \zeta(2m) - 1, \\ C_2 &= \sqrt{6}, & \epsilon_2(m) &= \sqrt{\frac{2\zeta(2m)^2 + \zeta(4m)}{3}} - 1, \\ C_\infty &= 4, & \epsilon_\infty(m) &= (1 - 2^{-2m})\zeta(2m) - 1. \end{aligned} \tag{11}$$

These are all easily verified independently of the general formulae.

The following is the Euler–MacLaurin summation formula in the form needed by the next section. The central idea is to integrate by parts repeatedly, choosing constants of integration so as to avoid differentiating the integrand in the interior of the domain of integration.

THEOREM 5. *If $f \in C^{2m}([0,1])$ then*

$$\int_0^1 f(t)\,dt = \frac{1}{2n}\sum_{j=0}^{n-1}\left(f\left(\frac{j}{n}\right)+f\left(\frac{j+1}{n}\right)\right)$$
$$-\sum_{k=1}^{m-1}\frac{B_{2k}^*(0)}{(2k)!}n^{-2k}\left(f^{(2k-1)}(1)-f^{(2k-1)}(0)\right)$$
$$+\frac{n^{-2m}}{(2m)!}\int_0^1\left(B_{2m}^*(nt)-B_{2m}^*(0)\right)f^{(2m)}(t)\,dt. \qquad (12)$$

Proof. If $g \in C^{2m}([0,1])$ then an integration by parts using (9) gives

$$\int_\epsilon^{1-\epsilon} g(x)\,dx = B_1^*(1-\epsilon) - B_1^*(\epsilon) - \int_\epsilon^{1-\epsilon} B_1^*(x)g'(x)\,dx.$$

A further $2m-2$ integration by parts using Eq. (10) gives

$$\int_\epsilon^{1-\epsilon} g(x)\,dx$$
$$=\sum_{k=1}^{2m-1}\frac{(-1)^{k+1}}{k!}\left[B_k^*(1-\epsilon)g^{(k-1)}(1-\epsilon)-B_k^*(\epsilon)g^{(k-1)}(\epsilon)\right]$$
$$-\int_\epsilon^{1-\epsilon}\frac{B_{2m-1}^*(x)}{(2m-1)!}g^{(2m-1)}(x)\,dx.$$

A final integration by parts, using

$$\frac{d}{dx}\left(B_{2m}^*(x)-B_{2m}^*(0)\right) = 2mB_{2m-1}^*(x),$$

gives

$$\int_\epsilon^{1-\epsilon} g(x)\,dx$$
$$=\sum_{k=1}^{2m-1}\frac{(-1)^{k+1}}{k!}\left[B_k^*(1-\epsilon)g^{(k-1)}(1-\epsilon)-B_k^*(\epsilon)g^{(k-1)}(\epsilon)\right]$$
$$+\frac{B_{2m}^*(1-\epsilon)-B_{2m}^*(0)}{(2m)!}f^{(2m-1)}(1-\epsilon)$$
$$-\frac{B_{2m}^*(\epsilon)-B_{2m}^*(0)}{(2m)!}f^{(2m-1)}(\epsilon)$$
$$+\int_\epsilon^{1-\epsilon}\frac{B_{2m}^*(x)-B_{2m}^*(0)}{(2m)!}g^{(2m)}(x)\,dx.$$

Taking limits as ϵ tends to zero, we obtain

$$\int_0^1 g(x) = \frac{1}{2}g(0) + \frac{1}{2}g(1) + \sum_{k=1}^{m-1} \frac{B_{2k}(0)}{(2k)!}\left(g^{(2k-1)}(1) - g^{(2k-1)}(0)\right)$$
$$+ \int_0^1 \frac{B^*_{2m}(x) - B^*_{2m}(0)}{(2m)!} g^{(2m)}(x)\,dx.$$

Making the substitution $x = nt - j$, $g(x) = n^{-1}f(t)$ and summing over j gives (12).

The integral on the right in the Euler−MacLaurin summation formula is an error term, to be estimated rather than evaluated. Using the norms calculated previously and Hölder's inequality, we find

$$\left| \int_0^1 \left(B^*_{2m}(nt) - B^*_{2m}(0)\right) f^{(2m)}(t)\,dt \right| \le K \|f^{(2m)}\|_{L^p([0,1])}, \tag{13}$$

where the constant K is given by

$$K = C_p\big(1 + \epsilon_p(m)\big)\frac{(2m)!}{(2\pi)^{2m}}.$$

The norms on the right are rarely known but can often be estimated, for example by Cauchy's theorem. If f is analytic and bounded on the set,

$$U_r = \left\{ z \in \mathbf{C} \colon \min_{t \in [0,1]} |z - t| < r \right\},$$

for example, then

$$\left| \int_0^1 \left(B^*_{2m}(nt) - B^*_{2m}(0)\right) f^{(2m)}(t)\,dt \right| \le K' \|f\|_{L^\infty(U_r)}, \tag{14}$$

with the constant K' given by

$$K' = 2\{\zeta(2m)\}\frac{(2m)!^2}{(2\pi r)^{2m}}.$$

Unlike the estimate (13), this one is not quite sharp. For complex analytic integrands an essentially different argument using the calculus of residues gives the error as a contour integral. Estimating this integral requires nontrivial information on the Binet function and some familiarity with Hardy spaces. In view of Theorem 15 this extra effort cannot be well rewarded, so we choose to avoid it.

The treatment of Bernoulli functions given here constitutes the tip of a rather interesting iceberg. The reader desiring a quick tour of the submerged portion is referred to [11].

6. NUMERICAL INTEGRATION

In this section we apply the extrapolation algorithm of Section 4 to a sequence of Riemann sums to obtain algorithms for numerical integration. Any proper definite integral of a function real analytic over a closed (finite) interval can be broken into a finite number of pieces to which the following theorem applies.

THEOREM 6. *If $f \in C^\infty([0,1])$ and if a_n is the Riemann sum,*

$$a_n = \frac{1}{2n} \sum_{j=0}^{n-1} \left(f\left(\frac{j}{n}\right) + f\left(\frac{j+1}{n}\right)\right) \tag{15}$$

corresponding to a regular mesh of size n^{-1}, then the sequence A_m defined by

$$A_m = \frac{2}{(2m)!} \sum_{n=1}^{m} (-1)^{m+n} \binom{2m}{m+n} n^{2m} a_n \tag{16}$$

satisfies the estimate,

$$\left| \int_0^1 f(t)\, dt - A_m \right| \le C_p \frac{\pi^{-2m}}{(2m)!} \| f^{(2m)} \|_{L^p([0,1])}. \tag{17}$$

If f is analytic and bounded in the set,

$$U_r = \left\{ z \in \mathbf{C}\colon \min_{t \in [0,1]} |z - t| < r \right\},$$

with $r > \pi^{-1}$ then A_m converges geometrically,

$$\left| \int_0^1 f(t)\, dt - A_m \right| < 2(\pi r)^{-2m} \| f \|_{L^\infty(U_r)}. \tag{18}$$

Proof. In view of the definition (2) of the $w^l_{m,n}$, Eq. (16) can be rewritten in the form,

$$A_m = \sum_{n=1}^{m} w^2_{m,n} a_n.$$

By Theorem 5 the sequence a_n satisfies the hypotheses of Theorem 3. The bound (13) on the error term gives the estimate,

$$\left| \int_0^1 f(t)\, dt - A_m \right| \leq K \| f^{(2m)} \|_{L^p([0,1])},$$

with

$$K = C_p \big(1 + \epsilon_p(m)\big) \frac{\pi^{-2m}}{(2m)!}.$$

In view of the remark following Theorem 3 this may be improved to

$$K = C_p \big(1 + \epsilon_p(m)\big) \left(1 - 2^{-2m} \binom{2m}{m}\right) \frac{\pi^{-2m}}{(2m)!}.$$

Because

$$\big(1 + \epsilon_p(m)\big) \left(1 - 2^{-2m} \binom{2m}{m}\right) < 1,$$

for all $1 \leq p \leq \infty$ and $m > 0$ we obtain 17. The statement about analytic functions is an immediate consequence of the Cauchy estimates,

$$\| f^{(2m)} \|_{L^\infty([0,1])} \leq (2m)!\, r^{2m} \| f \|_{L^\infty(U_r)}.$$

The auxiliary sequence A_m of the preceding theorem does not converge if f is merely analytic on $(0,1)$, and not in a neighborhood of the endpoints. This situation is typical, for example, when the integral is obtained by a change of variable from an integral over an infinite interval. See Section 12 for further discussion of this point. The following theorem is designed to deal with certain essential singularities. Essentially, it says that a function can still be integrated in polynomial time if it belongs to a Gevrey class.

THEOREM 7. *If $f \in C^\infty([0,1])$, and if a_n is the Riemann sum,*

$$a_n = \frac{1}{2n} \sum_{j=0}^{n-1} \left(f\left(\frac{j}{n}\right) + f\left(\frac{j+1}{n}\right) \right)$$

corresponding to a regular mesh of size n^{-1}, then the sequence A_m defined by

$$A_m = \frac{1}{(2rm)!} \sum_{n=1}^{rm} (-1)^{rm+n} \binom{2rm}{rm+n} n^{2rm} a_{dn^r}$$

satisfies the estimate,

$$\left| \int_0^1 f(t)\, dt - A_m \right| \leq K'' \| f^{(2m)} \|_{L^p([0,1])},$$

where

$$K'' = C_p\big(1 + \epsilon_p(m)\big)\frac{1}{(2rm)!}\left(\frac{2^{r-1}}{d\pi}\right)^{2m},$$

and C_p and ϵ_p as in the preceding section.

Proof. Similar to that of Theorem 6 with Theorem 4 in place of 3.

7. FINITE PRECISION

The preceding results all assume that the sequence a_n, whose limit is desired, is known to arbitrary precision. This assumption can be relaxed with very little effort, in order to get d digits of accuracy in the final result, we will need to know the original sequence to $O(d)$ digits of accuracy. The point of this section is to show that we need only $O(d)$ digits. The estimates of the effects of round-off error are hardest in the case $l = 2$ of Theorem 3, because that algorithm assigns the greatest weight to the early elements of the sequence. Here we treat only that case. The interested reader should have no difficulty in applying the same method to Theorem 4 and the case $l = 1$ of Theorem 3. We make no particular effort in this section to obtain sharp bounds.

Suppose, then, that we use

$$\tilde{a}_n = a_n + \delta_n,$$

in place of a_n. We will then obtain an auxiliary sequence,

$$\tilde{A}_m = A_m + \Delta_m,$$

where

$$\Delta_m = \frac{2}{(2m)!}\sum_{n=1}^{m}(-1)^{m+n}\binom{2m}{m+n}n^{2m}\delta_n. \tag{19}$$

We want to know how fast δ_n must decrease to make Δ_m decrease exponentially, and thus preserve the geometric convergence of A_m. The following lemma is not sharp, but is good enough for our present purposes.

LEMMA 8. *Let $l > 1$ and $n > 0$. Then*

$$\frac{1}{(2m)!}\binom{2m}{m+n}(ln)^{2m} < e^{2ln}, \tag{20}$$

for all $m > 0$.

Proof. Clearly

$$\frac{1}{(2m)!}\binom{2m}{m+n}(ln)^{2m} < \frac{1}{(2m)!}\binom{2m}{m}(ln)^{2m}.$$

Viewing the right-hand side as a function of m, and checking the ratio of successive terms, we see that the maximum occurs at $m = ln$, so

$$\frac{1}{(2m)!}\binom{2m}{m}(ln)^{2m} \leq \frac{(ln)^{2ln}}{(ln)!^2}.$$

Because all terms in the sum,

$$\sum_{j=0}^{\infty} \frac{(ln)^j}{j!} = e^{ln}$$

are positive, the sum must be greater than its lnth term,

$$\frac{(ln)^{ln}}{(ln)!} < e^{ln}.$$

Combining the three earlier inequalities gives (20).

The following theorem says that geometric decrease of $|\delta_n|$ implies that of $|\Delta_m|$.

THEOREM 9. *If* Δ_m *is defined by* (19), *and*

$$|\delta_n| \leq n^{-2}e^{-2ln},$$

then

$$|\Delta_m| < \frac{\pi^2}{3} l^{-2m}.$$

Proof. By the triangle inequality,

$$|\Delta_m| \leq \frac{2}{(2m)!} l^{-2m} \sum_{n=1}^{m} \left(\binom{2m}{m+n}(ln)^{2m}\right)|\delta_n|.$$

The preceding lemma shows that

$$|\Delta_m| < 2l^{-2m} \sum_{n=1}^{m} e^{2ln}|\delta_n| \leq 2l^{-2m} \sum_{n=1}^{m} n^{-2}.$$

Replacing the finite sum by the corresponding infinite sum gives the desired estimate.

8. ESTIMATES ON SUMS

The weights $w^l_{m,n}$ of Section 3 were chosen for simplicity rather than for performance. In Section 10 we introduce a new algorithm with weights $W^l_{m_s,m_d,n}$ which give faster convergence. The hardest part of the analysis of that algorithm is the analogue of Lemma 2. This requires an entertaining, but rather lengthy, digression into classical analysis and occupies this entire section. For the convenience of readers whose patience is limited, the main results are summarized in Lemma 10 at the end of the section. The properties of the gamma function used in the following text are all found in [2] or [7].

The new weights are

$$\begin{aligned} W^1_{m_d,m_s,n} &= \frac{n^m}{2^{m_s} m_d!}\Omega_{m_d,m_s,n}, \\ W^2_{m_d,m_s,n} &= \frac{2n^{2m}}{2^{2m_s}(2m_d)!}\Omega_{2m_d,2m_s,m+n}, \end{aligned} \tag{21}$$

where $m = m_d + m_s$, here and throughout this section, and

$$\Omega_{m_d,m_s,n} = \sum_{\substack{0 \le n_d \le m_d \\ 0 \le n_s \le m_s \\ n_d + n_s = n}} (-1)^{m_d+n_d} \binom{m_d}{n_d}\binom{m_s}{n_s}.$$

The sums to be estimated are

$$\sum_{n=1}^{m} |W^1_{m_d,m_s,n}| n^{-m} < \frac{1}{2^{m_s} m_d!} \sum_{n=0}^{m} |\Omega_{m_d,m_s,n}|,$$

and

$$\sum_{n=1}^{m} |W^2_{m_d,m_s,n}| n^{-2m} < \frac{1}{2^{2m_s}(2m_d)!} \sum_{n=0}^{2m} |\Omega_{2m_d,2m_s,m+n}|.$$

In proving the second estimate we make use of the obvious identity,

$$W^2_{m_d,m_s,n} = W^2_{m_d,m_s,-n}.$$

Thus it suffices to estimate sums of the form,

$$\sum_{n=0}^{m} |\Omega_{m_d,m_s,n}|.$$

Our main tool will be the Cauchy−Schwarz inequality. If α_n is an arbitrary sequence of nonnegative numbers then this inequality gives the estimate,

$$\begin{aligned}\sum_{n=0}^{m} |\Omega_{m_d, m_s, n}| &= \sum_{n=0}^{m} \alpha_n^{-1/2} \alpha_n^{1/2} |\Omega_{m_d, m_s, n}| \\ &\le \left(\sum_{n=0}^{m} \alpha_n^{-1} \right)^{1/2} \left(\sum_{n=0}^{m} \alpha_n \Omega_{m_d, m_s, n}^2 \right)^{1/2}.\end{aligned}$$

We wish to choose α_n so that both sums can be evaluated explicitly and so that their product is nearly minimal. Trial and error suggests the choice,

$$\alpha_n = (2n - m)^2 + m,$$

and we proceed to estimate the sums.

The first sum can be estimated using the Euler−MacLaurin summation formula from Section 5. Using the complex form (14) of the error estimate with $r = (2m)^{-1/2}$, and none of the endpoint derivatives, we obtain

$$\left| \sum_{n=0}^{m} \alpha_n^{-1} - \frac{1}{m^2 + m} - \frac{\arctan\sqrt{m}}{\sqrt{m}} \right| \le \frac{2}{3m},$$

from which we see that

$$\sum_{n=0}^{m} \alpha_n^{-1} \le \frac{\pi}{2\sqrt{m}},$$

at least for $m \ge 2$ and the case $m = 1$ is easily verified directly.

We now proceed to the second sum. The main idea here is to use the Plancherel theorem to replace the ℓ^2 norm of a sequence with the L^2 norm of a function. We define

$$\omega_{m_d, m_s}(t) = \sum_{n=0}^{m} \Omega_{m_d, m_s, n} e^{\pi i (2n - m) t},$$

and we compute

$$\left[m - \pi^{-2} \frac{d^2}{dt^2} \right] \omega_{m_d, m_s}(t) = \sum_{n=0}^{m} \alpha_n \Omega_{m_d, m_s, n} e^{\pi i (2n - m) t},$$

so Plancherel's theorem—or simply the distributive law, because all the sums are finite—gives the identity,

$$\sum_{n=0}^{m} \alpha_n \Omega_{m_d, m_s, n}^2 = \int_0^1 \omega_{m_d, m_s}(t) \left[m - \pi^{-2} \frac{d^2}{dt^2} \right] \omega_{m_d, m_s}(t)\, dt.$$

Clearly,

$$\omega_{m_d, m_s}(t) = 2^m i^{-m_d} \sin^{m_d} \pi t \cos^{m_s} \pi t,$$

so we can write the integrand as a polynomial in $\sin \pi t$ and $\cos \pi t$. This can be integrated using Euler's beta integral in the form,

$$2\int_0^{\pi/2} \sin^{2x-1} \theta \cos^{2y-1} \theta \, d\theta = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

Applying this to our integral, and using the functional equation,

$$\Gamma(z+1) = z\Gamma(z),$$

repeatedly, we obtain the formula,

$$\sum_{n=0}^{m} \alpha_n \Omega^2_{m_d, m_s, n} = 2^{2m+1}\pi^{-1} m\left(1 + \delta_{m_d, m_s}\right) \times \frac{\Gamma\left(m_d + \frac{1}{2}\right)^2 \Gamma\left(m_2 + \frac{1}{2}\right)}{\Gamma(m+1)},$$

where

$$\delta_{m_d, m_s} = \left(\frac{1}{8m_d - 4} + \frac{1}{8m_s - 4}\right)^{1/2} - 1.$$

Clearly δ_{m_d, m_s} tends to zero as m_d and m_s tend to infinity. We then use Legendre's duplication formula,

$$\Gamma(z)\Gamma\left(z + \tfrac{1}{2}\right) = 2^{1-2z}\pi^{1/2}\Gamma(2z),$$

and the functional equation to obtain the relation,

$$\frac{\Gamma\left(m_d + \frac{1}{2}\right)\Gamma\left(m_2 + \frac{1}{2}\right)}{\Gamma(m+1)} = 2^{2-2m}\pi m^{-1} \frac{\Gamma(2m_d)\Gamma(2m_s)}{\Gamma(m_d)\Gamma(m_s)\Gamma(m)}.$$

From Binet's second integral representation of the gamma function,

$$\Gamma(z) = \sqrt{\frac{2\pi}{z}} \left(\frac{z}{e}\right)^z \exp\left(2\int_0^\infty \frac{\arctan(t/z)\, dt}{e^{2\pi t} - 1}\right),$$

and the fact that the integral of a nonnegative function is nonnegative, we obtain the estimate,

$$\frac{\Gamma(2m_d)\Gamma(2m_s)}{\Gamma(m_d)\Gamma(m_s)\Gamma(m)} \leq 2^{2m-3/2}\pi^{-1/2} \frac{m_d^{m_d} m_s^{m_s}}{m^{m-1/2}}.$$

Combining the preceding estimates we find

$$\sum_{n}^{m} \alpha_n \Omega^2_{m_d, m_s, n} \leq 2^{2m+3/2} \pi^{-1/2} \left(1 + \delta_{m_d, m_s}\right)^2 \frac{m_d^{m_d} m_s^{m_s}}{m^{m-1/2}},$$

and hence,

$$\sum_{n}^{m} |\Omega_{m_d, m_s, n}| \leq \sqrt[4]{2\pi} \left(1 + \delta_{m_d, m_s}\right) \frac{2^m m_d^{m_d/2} m_s^{m_s/2}}{m^{m/2}}.$$

Applying this estimate to the desired sums gives

$$\sum_{n=1}^{m} |W^l_{m_d, m_s, n}| n^{-lm} \leq \frac{\sqrt[4]{2\pi}}{(lm_d)!} \left(1 + \delta_{lm_d, lm_s}\right) \frac{(4m_d)^{lm_d/2} m_s^{lm_s/2}}{m^{lm/2}}.$$

For given m_d, m_s this estimate is nearly sharp. In practice, however, we are given m and we are free to choose m_d, m_s subject to the constraint $m_d + m_s = m$. We will want to do so in such a way as to minimize the quantity,

$$(lm_d)! \rho^{-lm_d} \sum_{n=1}^{m} |W^l_{m_d, m_s, n}| n^{-lm},$$

with some $\rho > 0$. We have just seen that we can bound this by

$$\sqrt[4]{2\pi} \left(1 + \delta_{lm_d, lm_s}\right) \frac{(4\rho^{-2} m_d)^{lm_d/2} m_s^{lm_s/2}}{m^{lm/2}}.$$

Because the first factor is constant, and the second nearly so, we concentrate our efforts on the remaining factors. If we define

$$\mu_d = \frac{\rho^2}{\rho^2 + 4} m,$$

and

$$\mu_s = \frac{4}{\rho^2 + 4} m,$$

then

$$\mu_d + \mu_s = m,$$

and

$$\frac{(4\rho^{-2}\mu_d)^{l\mu_d/2} \mu_s^{l\mu_s/2}}{m^{lm/2}} = \left(1 + \frac{\rho^2}{4}\right)^{-lm/2}.$$

Usually μ_d and μ_s are not integers, so we take m_d and m_s to be the integers closest to μ_d and μ_s, and we note that μ_d and μ_s both tend to infinity as m does. We compute

$$\left(1+\delta_{lm_d, lm_s}\right)\frac{\left(4\rho^{-2}\mu_d\right)^{l\mu_d/2}\mu_s^{l\mu_s/2}}{m^{lm/2}} = \left(1+\varepsilon_{l,\rho,m}\right)\left(1+\frac{\rho^2}{4}\right)^{-lm/2},$$

where

$$\varepsilon_{l,\rho,m} = \left(1+\delta_{lm_d, lm_s}\right)\left\{\left(1+\frac{\mu_d-m_d}{m_d}\right)^{-lm_d/2} e^{l(\mu_d-m_d)/2}\right\}$$
$$\times\left\{\left(1+\frac{m_s-\mu_s}{m_s}\right)^{-lm_s/2} e^{l(m_s-\mu_s)/2}\right\} - 1.$$

Note that the relation,

$$\mu_d+\mu_s = m_d+m_s$$

causes the exponentials to cancel. We also note that, as m tends to infinity, each of the three factors in the foregoing text tends to 1, the first obviously and the second and third because of Euler's formula,

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^{-n} e^x = 1.$$

We summarize our results in the following lemma.

LEMMA 10. *Let* $l = 1$ *or* 2, *and define* $W^l_{m_d, m_s, n}$ *by* (21). *The following estimate holds*

$$\sum_{n=1}^{m} |W^l_{m_d,m_s,n}| n^{-lm} \le \frac{\sqrt[4]{2\pi}}{(lm_d)!}\left(1+\delta_{lm_d, lm_s}\right)\frac{(4m_d)^{lm_d/2} m_s^{lm_s/2}}{m^{lm/2}}.$$

If $\rho > 0$ *and* $m > 0$ *are given we can choose* m_d *and* m_s *such that* $m_d + m_s = m$ *and*

$$(lm_d)!\,\rho^{-lm_d}\sum_{n=1}^{m} |W^l_{m_d,m_s,n}| n^{-lm} \le \sqrt[4]{2\pi}\left(1+\varepsilon_{l,\rho,m}\right)\left(1+\frac{\rho^2}{4}\right)^{-lm/2}.$$

The quantities δ_{lm_s, lm_d} *and* $\varepsilon_{l,\rho,m}$ *are small in the sense that*

$$\lim_{\substack{m_d \to \infty \\ m_s \to \infty}} \delta_{lm_s, lm_d} = 0,$$

and

$$\lim_{m \to \infty} \varepsilon_{l,\rho,m} = 0.$$

9. SMOOTHING OPERATORS

The extrapolation algorithm of Section 4 is based on the properties of the difference operators of Section 3. These are the simplest approximations to differentiation, but they are not the best for our present purposes. The rate of convergence of the extrapolation algorithms can be improved by composing these difference operators with certain discrete smoothing operators. This section is devoted to proving Lemmas 11 and 12, which generalize Lemmas 1 and 2 of Section 3.

If the operators S_1, S_2 on $\mathbf{C}[x]$ are defined by

$$(S_1 p)(x) = p(x+1) + p(x),$$
$$(S_2 p)(x) = p\left(x + \tfrac{1}{2}\right) + p\left(x - \tfrac{1}{2}\right),$$

then the following formulae are easily proved by induction on m,

$$(S_1^m p)(x) = \sum_{n=0}^{m} \binom{m}{n} p(x+n),$$
$$\left(S_2^{2m} p\right)(x) = \sum_{n=-m}^{m} \binom{2m}{m+n} p(x+n),$$
$$S_l^m x^k = 2^m x^k + \text{lower order terms}.$$

Combining these with the corresponding formulae for D, and evaluating at $x = 0$, gives formulae which can be written concisely in terms of the $W^l_{m_d, m_s, n}$ of (21) as

$$\sum_{n=1}^{m} W^l_{m_d, m_s, n} n^{-lk} = \begin{cases} 1, \text{if } k = 0, \\ 0, \text{if } 0 < k < m_d, \\ ?, \text{otherwise}, \end{cases}$$

where $m = m_d + m_s$ here and throughout this section. This formula is needed in the proof of

LEMMA 11. *Let $l = 1$ or 2. If*

$$p_n = \sum_{k=0}^{m_d - 1} c_k n^{-lk},$$

and

$$P_{m_d, m_s} = \sum_{n=1}^{m} W^l_{m_d, m_s, n} p_n,$$

with $W^l_{m_d, m_s, n}$ given by (21), *then $P_{m_d, m_s} = c_0$.*

Proof. Bring the summation with respect to k outside the summation with respect to n, and use the formula derived previously.

LEMMA 12. *Let $l = 1$ or 2. If*

$$|r_n| \le e_{m_d} n^{-lm_d},$$

and

$$R_{m_d, m_s} = \sum_{n=1}^{m} W^l_{m_d, m_s, n} r_n,$$

with $W^l_{m_d, m_s, n}$ given by (21), *then*

$$|R_{m_d, m_s}| \le \frac{\sqrt[4]{2\pi}}{(lm_d)!}\left(1 - \delta_{lm_d, lm_s}\right)\left(\frac{4m_d}{m}\right)^{lm_d/2}\left(\frac{m_s}{m}\right)^{lm_s/2},$$

where δ_{lm_d, lm_s} tends to zero for m_d and m_s large. If $\rho > 0$ and $m > 0$ are given such that

$$e_\mu \le (l\mu)!\rho^{-l\mu},$$

for all $\mu \le m$, then it is possible to choose integers m_d, m_s such that

$$m_d + m_s = m,$$

and

$$|R_{m_d, m_s}| \le \sqrt[4]{2\pi}\left(1 + \varepsilon_{l, \rho, m}\right)\left(1 + \frac{\rho^2}{4}\right)^{-lm/2},$$

with $\varepsilon_{l, \rho, m}$ such that

$$\lim_{m \to \infty} \varepsilon_{l, \rho, m} = 0.$$

Proof. We have

$$|R_{m_d,m_s}| \leq \max_{1 \leq n \leq m} |r_n n^{lm}| \sum_{n=1}^{m} |W^l_{m_d,m_s,n} n^{-lm}|.$$

The first term is bounded by e_m by hypothesis. The second term has been estimated in Lemma 10.

10. IMPROVED ALGORITHMS

The smoothing operators of the preceding section allow us to improve the algorithms of Section 4. The main result of this section, Theorem 13, drops the hypothesis $\rho > 2$ from Theorem 3, while improving the rate of convergence in its conclusion.

THEOREM 13. *Let $l = 1$ or 2. If*

$$\left| a_n - \sum_{k=0}^{m_d - 1} c_k n^{-lk} \right| \leq e_{m_d} n^{-lm_d}, \tag{22}$$

and

$$A_{m_d,m_s} = \sum_{n=1}^{m} W^l_{m_d,m_s,n} a_n, \tag{23}$$

with $W^l_{m_d,m_s,n}$ given by (21), then

$$|A_{m_d,m_s} - c_0| \leq \frac{\sqrt[4]{2\pi}\, e_{m_d}}{(lm_d)!} \left(1 + \delta_{lm_d,lm_s}\right) \left(\frac{4m_d}{m}\right)^{lm_d/2} \left(\frac{m_s}{m}\right)^{lm_s/2},$$

where δ_{lm_d,lm_s} tends to zero for m_d and m_s large. If there are $C > 0$ and $\rho > 0$ such that

$$e_\mu \leq C(l\mu)!\rho^{-l\mu},$$

for all $\mu \leq m$, then we may choose m_d, m_s such that

$$m_d + m_s = m,$$

and

$$|A_{m_d,m_s} - c_0| \leq C\sqrt[4]{2\pi}\left(1 + \varepsilon_{l,\rho,m}\right)\left(1 + \frac{\rho^2}{4}\right)^{-lm/2}, \tag{24}$$

where $\varepsilon_{l,\rho,m}$ tends to zero for m large.

Proof. We define

$$\begin{aligned} p_n &= \sum_{k=0}^{m-1} c_k n^{-lk}, \\ P_{m_d, m_s} &= \sum_{n=1}^{m} W^l_{m_d, m_s, n} p_n, \\ r_n &= a_n - p_n, \\ R_{m_d, m_s} &= \sum_{n=1}^{m} W^l_{m_d, m_s, n} r_n. \end{aligned}$$

Clearly $R_{m_d, m_s} = A_{m_d, m_s} - P_{m_d, m_s}$. Lemma 11 shows that

$$P_{m_d, m_s} = c_0,$$

and Lemma 12 gives the estimates,

$$|R_{m_d, m_s}| \leq \frac{\sqrt[4]{2\pi}\, e_{m_d}}{(lm_d)!}\left(1 + \delta_{lm_d, lm_s}\right)\left(\frac{4m_d}{m}\right)^{lm_d/2}\left(\frac{m_s}{m}\right)^{lm_s/2},$$

and

$$|R_{m_d, m_s}| \leq C\sqrt[4]{2\pi}\left(1 + \varepsilon_{l, \rho, m}\right)\left(1 + \frac{\rho^2}{4}\right)^{-lm/2},$$

when m_d and m_s are the integers nearest $\rho^2 m/(\rho^2 + 4)$ and $4m/(\rho^2 + 4)$. For the benefit of any readers who need completely explicit estimates we state without proof the following bounds,

$$\begin{aligned} \delta_{lm_d, lm_s} &\leq \frac{1}{4m_d} + \frac{1}{4m_s}, \\ \varepsilon_{l, \rho, m} &\leq \frac{1}{m_d} + \frac{1}{m_s}. \end{aligned}$$

The same techniques could be used to prove an analogue of Theorem 4, but we will refrain from doing so. An improved extrapolation algorithm gives an improved algorithm for numerical integration.

THEOREM 14. *If $f \in C^\infty([0,1])$ and if a_n is the Riemann sum,*

$$a_n = \frac{1}{2n} \sum_{j=0}^{n-1} \left(f\left(\frac{j}{n}\right) + f\left(\frac{j+1}{n}\right)\right) \tag{25}$$

corresponding to a regular mesh of size n^{-1}, then the sequence A_{m_d, m_s} defined by

$$A_m = \sum_{n=1}^{m} W^2_{m_d, m_s, n} a_n, \tag{26}$$

with $W^2_{m_d, m_s, n}$ given by (21) *satisfies the estimate,*

$$\left| \int_0^1 f(t)\, dt - A_{m_d, m_s} \right| \leq \frac{2^{-7/4}\pi^{9/4}}{(2m_d)!} \left(\frac{m_d}{\pi^2 m} \right)^{m_d} \left(\frac{m_s}{m} \right)^{m_s} \| f^{(2m_d)} \|_{L^p([0,1])}. \tag{27}$$

If f is analytic and bounded in the set,

$$U_r = \left\{ z \in \mathbf{C}\colon \min_{t \in [0,1]} |z - t| < r \right\},$$

then we can choose, for each m, m_d, and m_s such that A_{m_d, m_s} converges geometrically,

$$\left| \int_0^1 f(t)\, dt - A_{m_d, m_s} \right| < 2^{1/4}\pi^{9/4}(1 + \pi^2 r^2)^{-m} \| f \|_{L^\infty(U_r)}. \tag{28}$$

Proof. Similar to that of Theorem 6 with Theorem 13 in place of Theorem 3. Here m_d and m_s are the integers nearest

$$\frac{\pi^2 r^2 m}{\pi^2 r^2 + 1} \quad \text{and} \quad \frac{m}{\pi^2 r^2 + 1}.$$

11. SHARPNESS OF RESULTS

In Section 4 we saw that if the basic estimate,

$$\left| a_n - \sum_{k=0}^{m-1} c_k n^{-k} \right| \leq C m! (\rho n)^{-m} \tag{29}$$

is satisfied for all m and n, and the time to compute a_n is $O(n^r)$ for some $r > 0$, then the limit c_0 can be computed to d digits in time $O(d^{r+1})$. The purpose of this section is to show that no algorithm gives time $O(d^{r+1-\epsilon})$ with $\epsilon > 0$. The algorithm of Section 4 can be improved, as shown in Section 10, but it cannot be improved by much. This is true even if we admit nonlinear methods.

The main result of this section, Theorem 15, is a theorem about theorems. It says that any theorem of the same general form as Theorems 3 and 13 must satisfy certain conditions on the auxiliary sequence A_m. What does the phrase "of the same general form as Theorems 3 and 13" mean? We assume that certain functions $A_{M,\rho}(a_1, a_2, \dots)$ are given for every $M \in \mathbf{N}$ and $\rho > 0$ and certain constants $K_\rho > 0$, $s_\rho > 0$ such that the basic estimate (29) implies the estimate,

$$\left| A_{M,\rho}(a_1, a_2, \dots) - c_0 \right| \le C K_\rho s_\rho^{-M}, \tag{30}$$

for all M. The functions and constants are allowed to depend on the values of their subscripts, but not on anything else. Note that, not only are the $A_{M,\rho}$ not assumed to depend linearly on their arguments, but no regularity of any kind is assumed. If, for some ρ, $s_\rho > 1$, then such a theorem gives an exponentially convergent series of approximants to c_0. What we wish to show is that, in this case, $A_{M,\rho}$ depends either on many of the a_n, or on at least one with n large. To this end we introduce the following set. Thus,

$$\Sigma_{M,\rho} = \left\{ n \in \mathbf{N} \colon A_{M,\rho} \text{ depends on the value of } a_n \right\}.$$

Thus we wish to show that either the cardinality $|\Sigma_{M,\rho}|$ or the supremum $\sup \Sigma_{M,\rho}$ is large. This is accomplished by the following theorem.

THEOREM 15. *Let r, $A_{M,\rho}$, K_ρ, s_ρ, and $\Sigma_{M,\rho}$ be as in the previous text, i.e., such that the estimate* (29) *implies the estimate* (30). *If there is an ρ for which $s_\rho > 1$ then, for this ρ and all sufficiently large M, either*

1.

$$|\Sigma_{M,\rho}| > \frac{\log s_\rho}{2r+2} \frac{M}{\log M},$$

or

2.

$$\sup \Sigma_{M,\rho} > M^{r+1}.$$

Proof. Suppose, on the contrary, that

$$\sup \Sigma_{M,\rho} < M^{r+1},$$

and

$$\sup \Sigma_{M,\rho} < M^{r+1},$$

for some

$$M > 2\frac{\rho + \log K_\rho + 1}{\log s_\rho}.$$

Define sequences,

$$a_n = \prod_{j \in \Sigma_{M,\rho}} \left(1 - \frac{j}{n}\right),$$

$$c_k = (-1)^k \sigma_k(\Sigma_{M,\rho}),$$

where σ_k is the kth elementary symmetric function. By construction,

$$a_n = \sum_{k=0}^{\infty} c_k n^{-k},$$

and

$$a_j = 0,$$

for all $j \in \Sigma_{M,\rho}$. In view of the definition of $\Sigma_{M,\rho}$ the latter property implies

$$A_{M,\rho}(a_1, a_2, \ldots) = A_{M,\rho}(0, 0, \ldots),$$

and the hypothesized theorem clearly shows that the quantity on the right must be 0.

Easy estimates give

$$\begin{aligned} n^m \left| a_n - \sum_{k=0}^{m-1} c_k n^{-k} \right| &= \left| \sum_{k=m}^{\infty} c_k n^{m-k} \right| \\ &\le \sum_{k=m}^{\infty} |c_k| \\ &\le \sum_{k=0}^{\infty} |c_k| \\ &= \prod_{j \in \Sigma_{M,\rho}} (1 + j) \end{aligned}$$

$$
\begin{aligned}
&\leq \left(\sup \Sigma_{M,\rho}\right)^{|\Sigma_{M,\rho}|} \\
&\leq M^{M \log s_\rho / 2 \log M} \\
&= s_\rho^{M/2} \\
&\leq m!\rho^{-m} e^{\rho} s_\rho^{M/2} \\
&\leq \frac{1}{eK_\rho} m!\rho^{-m} s_\rho^{M}.
\end{aligned}
$$

Thus the estimate (29) is satisfied with

$$C = \frac{1}{eK_\rho} s_\rho^M.$$

It now follows from the estimate (30) that

$$\left| A_{M,\rho}(a_1, a_2, \dots) - c_0 \right| \leq e^{-1}.$$

But we have already seen that the quantity on the left is equal to 1, so we have the desired contradiction.

12. MORE COMPLICATED INTEGRALS

Numerical integration of smooth functions over finite intervals is intended as an example of the general technique. Any serious application is likely to require some customization, and the present section is intended to give some general advice as to how this may be accomplished.

Integrals arising in applications generally differ from those considered here in three respects. The number of dimensions is generally greater than 1. The domain of integration is typically infinite. In addition, the integrands usually possess some singularities, at least on the boundary and often in the interior of the domain of integration. We discuss these in turn.

Fubini's theorem from elementary multivariable calculus shows that we can always reduce the evaluation of a multiple integral to that of several single integrals. The simplest example is that of the unit square,

$$R = \left\{(x, y) \in \mathbf{R}^2 : 0 \leq x, y \leq 1\right\},$$

where

$$\iint_R f(x, y)\, dA = \int_0^1 \left[\int_0^1 f(x, y)\, dx \right] dy.$$

If f is real analytic on R then clearly,

$$g_y(x) = f(x, y)$$

is an analytic function on $[0,1]$ for each fixed y, so the inner integral may be evaluated using Theorem 6. It is not difficult to show that

$$h(y) = \int_0^1 f(x, y)\, dx$$

is also real analytic on $[0,1]$, so the outer integral can also be evaluated using Theorem 6. Even though the geometry of this example is in some sense ideal for such a reduction there is a better approach. From the Riemann sum definition of the multiple integral it is clear that

$$\iint_R f(x, y)\, dA = \lim_{n\to\infty} n^{-2} \sum_{(j/n,\, k/n)\in R} f\left(\frac{j}{n}, \frac{k}{n}\right).$$

The same techniques used in Section 5 can be adapted to show that the sequence on the right possesses an asymptotic expansion in n^{-1} with error constants e_m which grow only factorially in m. We may therefore apply part (a) of Theorem 3 to accelerate the convergence of the sequence. This gives a significantly faster algorithm than the iterated integral.

Similar remarks apply to more complicated domains, but considerable care must be taken with the choice of Riemann sums. For example, let T be the triangle,

$$T = \{(x, y) \in \mathbf{R}^2 \colon x, y \geq 0,\ x + y \leq 1\},$$

and D the disc,

$$D = \{(x, y) \in \mathbf{R}^2 \colon x^2 + y^2 \leq 1\}.$$

We may write integrals over either domain as iterated integrals,

$$\iint_T f(x, y)\, dA = \int_0^1 \int_0^1 (1 - v) f((1 - v)u, v)\, du\, dv,$$

$$\iint_D f(x, y)\, dA = 2\pi \int_0^1 \int_0^1 f(u\cos(2\pi v), u\sin(2\pi v)u)\, du\, dv,$$

or as a limit of Riemann sums,

$$\iint_T f(x, y)\, dA = \lim_{n\to\infty} n^{-2} \sum_{(j/n,\, k/n)\in T} f\left(\frac{j}{n}, \frac{k}{n}\right),$$

$$\iint_D f(x, y)\, dA = \lim_{n\to\infty} n^{-2} \sum_{(j/n,\, k/n)\in D} f\left(\frac{j}{n}, \frac{k}{n}\right).$$

In either case we may use part (a) of Theorem 3 to evaluate the inner and outer integrals. In the case of T, it can be shown that the sequence of Riemann sums given earlier possesses an asymptotic expansion in n^{-1} similar to the Euler−MacLaurin formula, and that the extrapolation algorithm of Theorem 3 applied to this sequence gives better results than the iterated integral. Although the sequence given for D looks identical, one can show that it does *not* possess any asymptotic expansion, and the extrapolation algorithm of Theorem 3 gives garbage if applied to this sequence.

We turn now to integrals over infinite domains. Again, the examination of a special case is instructive. Consider the Laplace transform,

$$F(s) = \int_0^\infty f(t) e^{-st}\, dt,$$

of a function f analytic $[0, \infty)$. We assume that

$$|f(t)| \le Ce^{-s_0 t},$$

for all $t > 0$ with some $C > 0$ and $s_0 \in \mathbf{R}$, so that the integral defining $F(s)$ converges for $\operatorname{Re} s > s_0$. There are several approaches available for computing F. The definition of the infinite integral,

$$F(s) = \lim_{T \to \infty} \int_0^T f(t) e^{-st}\, dt$$

suggests approximating the integral over the infinite interval by an integral over a large finite interval, to which we can apply the results of Section 6. This will require estimates on the original function f in some neighborhood of the interval $[0, \infty)$ in the complex plane. In most applications f will satisfy estimates of the form,

$$|f(t)| \le C_\theta e^{-s_\theta \operatorname{Re} t}, \tag{31}$$

for all t with $-\theta \le \arg t \le \theta$ and for some $\theta > 0$, $C_\theta > 0$, $s_\theta < s$. Under these hypotheses we can use the algorithms of Section 6 to evaluate $F(s)$ to d digits of accuracy in time polynomial in d.

Alternatively we may make the preliminary change of variable $t = 1/u$ to transform the infinite interval $[1, \infty)$ into the finite interval $(0, 1]$,

$$F(s) = \int_0^1 f(t) e^{-st}\, dt + \int_0^1 u^{-2} f\left(\frac{1}{u}\right) e^{-s/u}\, du.$$

The new integrand is not analytic. Under the hypothesis (31) it is, however, of Gevrey class, and can be integrated using Theorem 7. In fact, this observation was one of our main reasons for stating Theorem 7. Again, the conclusion is that we can evaluate $F(s)$ to d digits of accuracy in time polynomial in d.

Finally we may approximate the infinite integral by an infinite Riemann sum,

$$F(s) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{\infty} f\left(\frac{j}{n}\right) e^{-js/n}.$$

This sum is sometimes known as a z-transform. The same techniques used to derive the Euler−MacLaurin expansion may be used to derive the asymptotic expansion of this sequence. The condition (31) guarantees that the boundary terms at infinity all vanish, and that the error constants do not grow too rapidly. Only in rare cases may the infinite sum be evaluated, but the definition of the infinite sum,

$$\sum_{j=0}^{\infty} f\left(\frac{j}{n}\right) e^{-js/n} = \lim_{N\to\infty} \sum_{j=0}^{N} f\left(\frac{j}{n}\right) e^{-js/n},$$

shows that we may approximate it arbitrarily closely by a finite sum, and Theorem 9 may be used to determine which N is sufficiently large that the truncation algorithm does not destroy the convergence. The conclusion is once again that we can evaluate $F(s)$ to d digits of accuracy in time polynomial in d.

Similar remarks apply to the third source of trouble, singularities of the integrand. In fact the real problem with the Laplace transform was that the integrand had an essential singularity at one of the endpoints, not that that endpoint happened to be at infinity. There are, however, several new observations worth making. Oscillatory integrals require special techniques. The main tool here is the method of stationary phase. Principal value integrals, by contrast, are easily handled. The general philosophy is to make the same cancellation occur in the Riemann sum as in the integral. Perhaps an elementary example will clarify this point. Suppose that k is real analytic on $[0,2]$ except for a simple pole of residue r at 1. We wish to compute the Cauchy principal value,

$$PV\int_0^2 k(t)\,dt = \lim_{\epsilon\to\infty}\left[\int_0^{1-\epsilon} k(t)\,dt + \int_{1+\epsilon}^2 k(t)\,dt\right].$$

By hypothesis,

$$k(t) = \frac{r}{t} + f(t),$$

where f is real analytic on $[0, 2]$. Consider the Riemann sum,

$$a_n = \frac{1}{2n} \sum_{j=0}^{2n-1} \left[k\left(\frac{j}{n}\right) + k\left(\frac{j+1}{n}\right) \right],$$

where we arbitrarily assign $k(0)$ the value 0. Clearly,

$$a_n = \frac{1}{2n} \sum_{j=0}^{2n-1} \left[f\left(\frac{j}{n}\right) + f\left(\frac{j+1}{n}\right) \right] - f(0) n^{-1},$$

while

$$PV \int_0^2 k(t)\, dt = \int_0^2 f(t)\, dt.$$

We may therefore apply the Euler–MacLaurin theorem to f to show that a_n has an asymptotic expansion in n^{-1} with constant term $PV \int_0^2 k(t)\, dt$, and Theorem 3 is applicable. Note that we do not need the residue r or the function f in order to evaluate the principal value integral by this method. It suffices to know they exist. A modification of this technique is applicable to all the singular integrals of Calderón–Zygmund type.

REFERENCES

1. A. C. Aitken, On Bernoulli's numerical solution of algebraic equations. *Proc. Roy. Soc. Edinburgh* **46** (1926), 289–305.
2. J. P. M. Binet, *J. Ecole Poly.* **XVI** (1839), 123–143.
3. R. Bulirsch and J. Stoer, "Introduction to Numerical Analysis," Springer-Verlag, New York, 1983.
4. B. Cavalieri, "Centuria di Varii Problemi," Bologna, p. 446, 1639.
5. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Numerical Recipes in C," Cambridge Univ. Press, Cambridge, U.K., 1992.
6. L. Euler, Methodus generalis summandi progressiones. *Comm. ASP* **VI** (1738), 68–97, reprinted *in* "Opera Omnia," Teuber, Leipzig, 1921.
7. L. Euler, *Nov. Comm. Petrop.* **XVI** (1772).
8. L. Euler, "Opera Omnia," Teubner, Leipzig, 1921.
9. C. G. J. Jacobi, De usu legitimo formulae summatoriae MacLaurinianae, *Crelle* **12** (1834), 263–272, reprinted *in* "Mathematische Werke," Reimer, Berlin, 1881.
10. C. G. J. Jacobi, "Mathematische Werke," Reimer, Berlin, 1881.
11. M. Lerch, Note sur la fonction $\sum_0^\infty e^{2\pi i x}/(w+k)^s$, *Acta Math.* **11** (1887–1888), 19–24.
12. C. MacLaurin, "Treatise on Fluxions," Ruddimans, Edinburgh, 1742.
13. L. F. Richardson, The deferred approach to the limit, *Philos. Trans. Roy. Soc. London* **226** (1926–1927), 299–349.
14. W. Romberg, Vereinfachte numerische Integration, *Det. Kgl. Norske Vid. Selsk. Forh.* **28** (1955), 68–97.